

# Analysis of Revisions to Aggregate Labor Productivity Measures

John Glaser, Kendra Hathaway Asher, Peter B. Meyer, Jay Stewart, and Jerin Varghese  
Office of Productivity and Technology, Bureau of Labor Statistics



We examine revisions to our headline quarterly Nonfarm Business Productivity growth figure.

The underlying statistic is approximately  $\frac{\text{output this quarter}}{\text{output last quarter}} \frac{\text{hours worked this quarter}}{\text{hours worked last quarter}}$   
(Actually it's  $\ln\left(\frac{\text{output per hour index this quarter}}{\text{output per hour index last quarter}}\right)$ .) Nonfarm business covers  $\frac{3}{4}$  of the U.S. economy.

Month	PFEI Releases	Revision notes
Feb.	Prelim for Q4, R2 for Q3	
March	R1 for Q4	Incorporates the annual CES benchmark revision through Q4 of the previous year. This affects both current and prior quarter hours.
May	Prelim for Q1, R2 for Q4	
June	R1 for Q1	
Aug.	Prelim for Q2, R2 for Q1	Both prelim and R2 estimates incorporate the annual NIPA/GDP benchmark revision (current and prior quarter output).
Sept.	R1 for Q2	
Nov.	Prelim for Q3, R2 for Q2	
Dec.	R1 for Q3	

## Calendar of releases

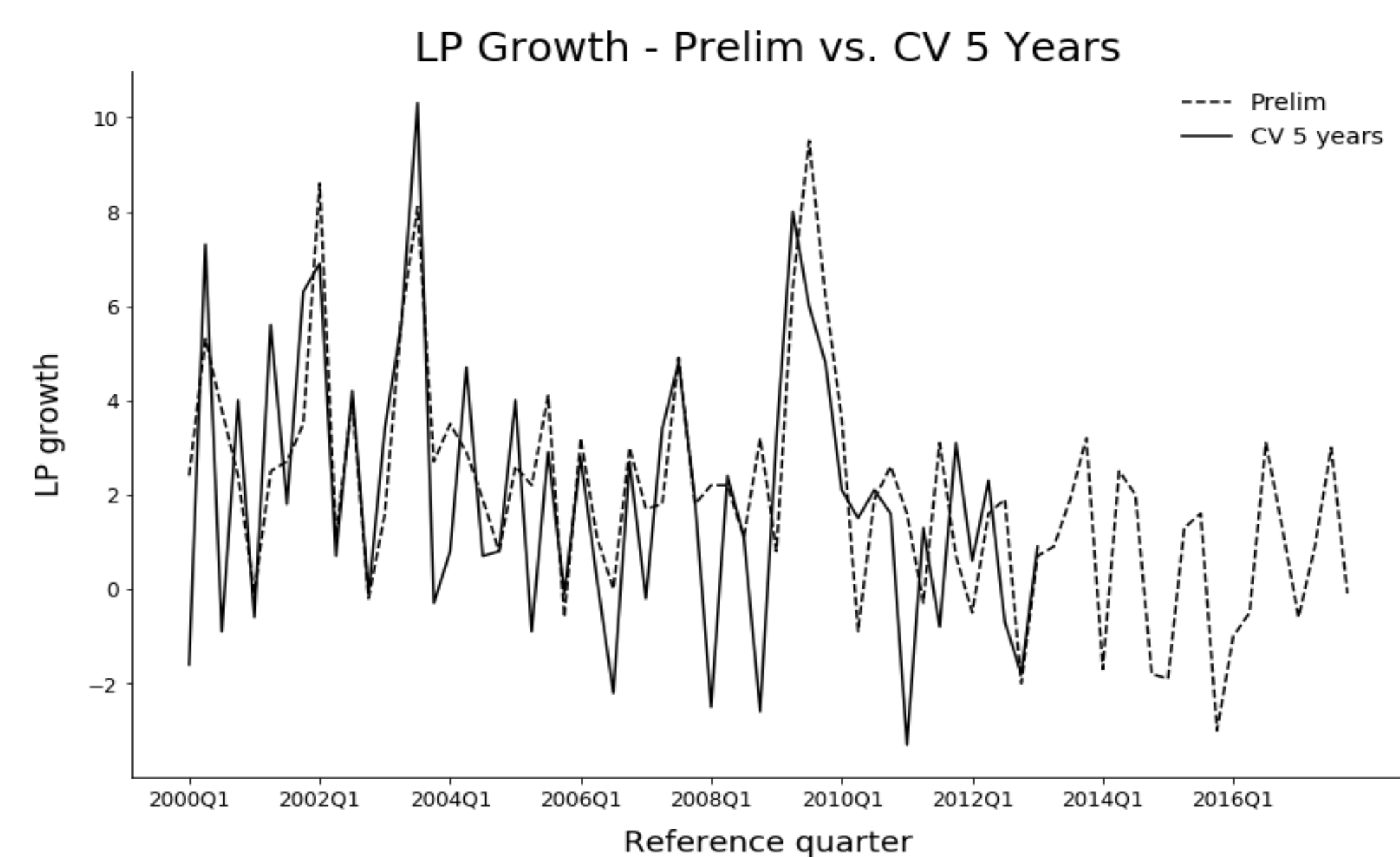
We issue the first three estimates in news releases. There are two releases per quarter. At each release, all previous quarters are re-estimated. We examine vintages of data for the same reference quarter:

- Preliminary (prelim), released ~40 days after quarter
- First revision (R1), released ~30 days after prelim
- Second revision (R2), released ~60 days after R1
- Current value (CV), latest estimate, including NIPA Comprehensive Revisions to methods.

## Data here

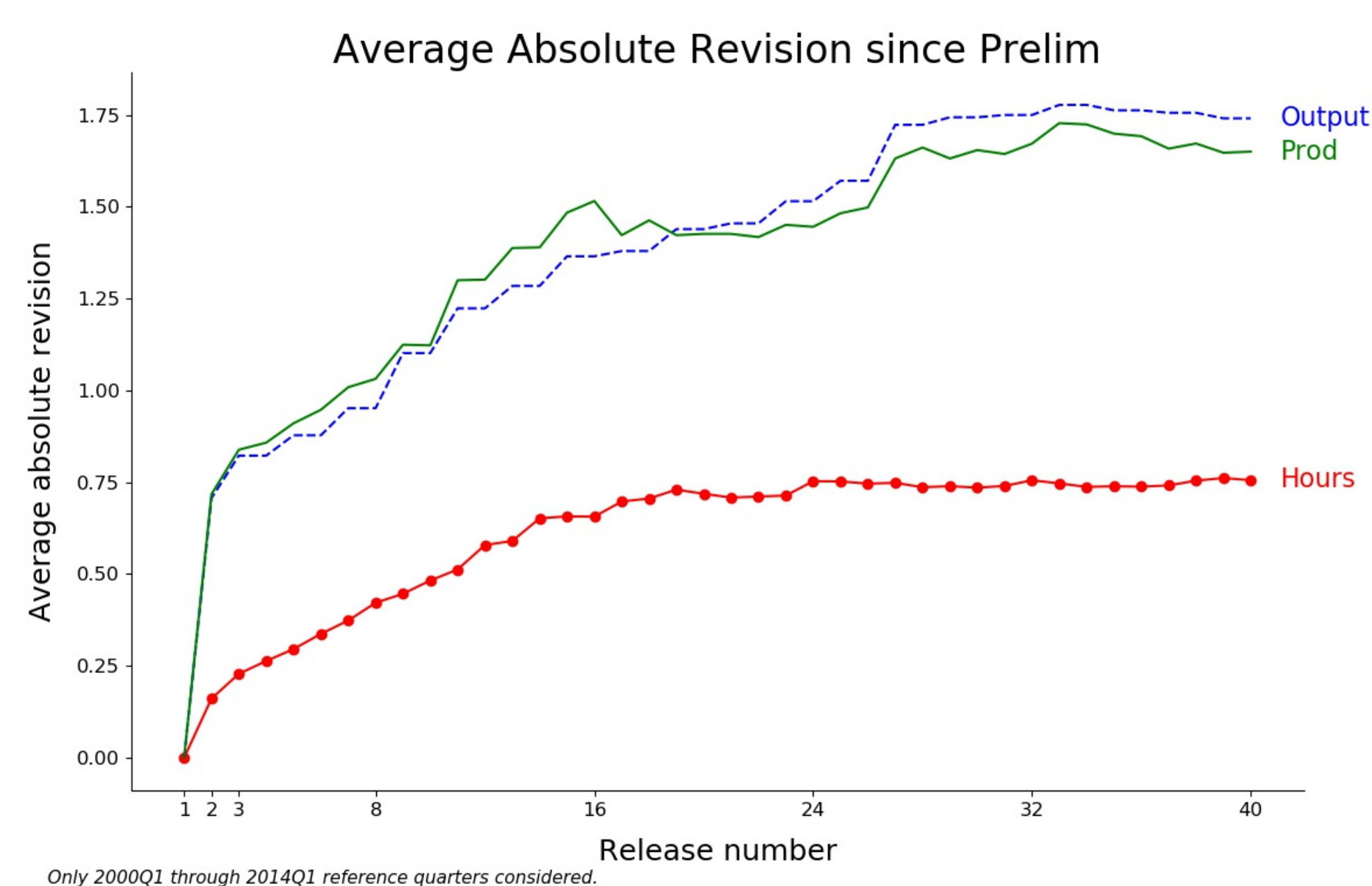
Output, hours, and productivity growth for 2000Q1-2014Q1, and 5 years of revisions. All are seasonally adjusted and annualized.

← Quarterly labor productivity (LP) growth is revised substantially. Revisions rarely change the overall story except for the 2008 recession period.



## Sources of revisions to hours and output

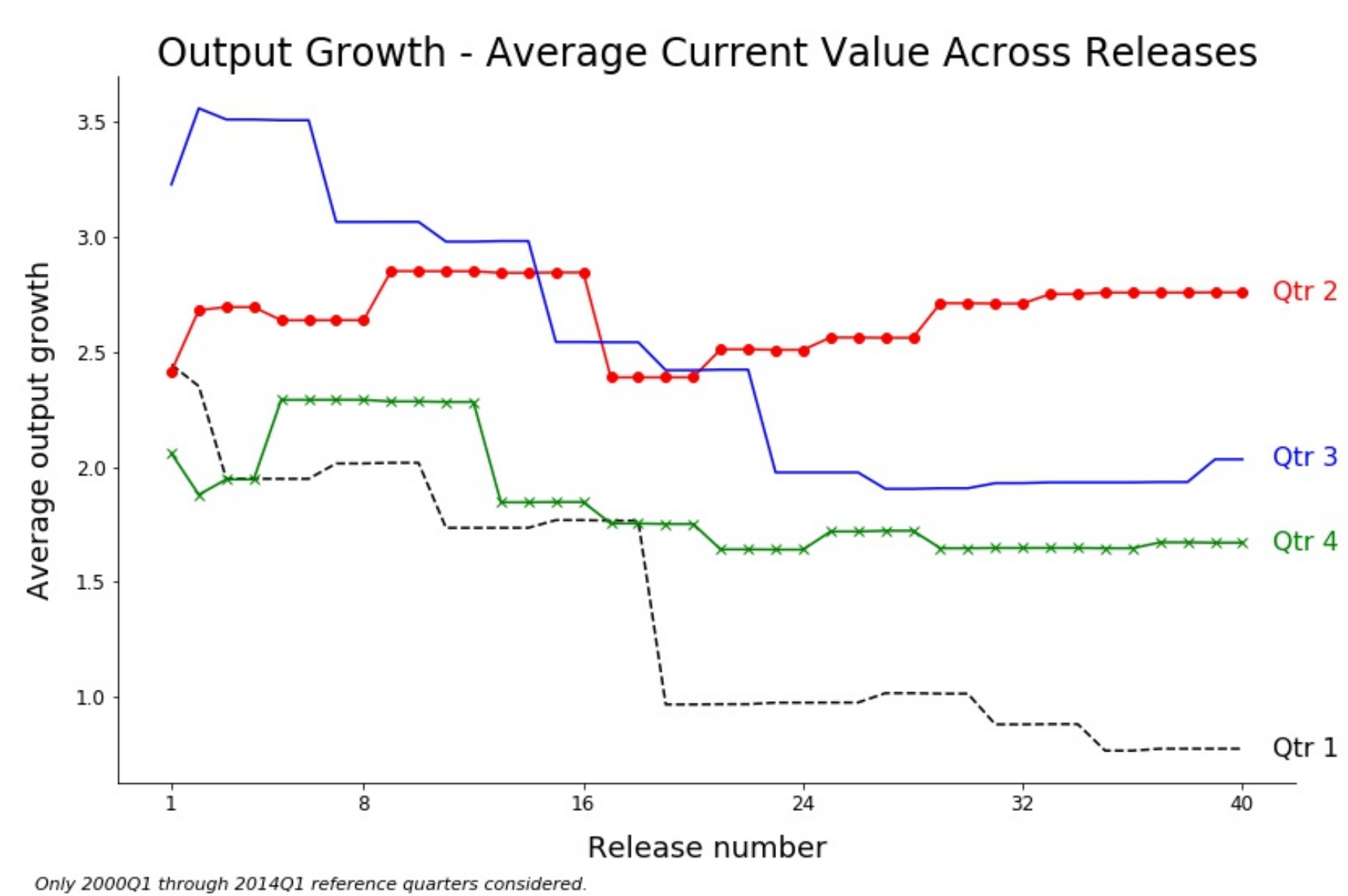
New microdata, seasonal adjustment, benchmarking to other data, comprehensive revisions of methods, price index updates, data corrections -- in either current or prior quarter.



## Revisions stabilize after 2-3 years

Output revisions are more than twice as large as hours revisions. Underlying output growth is also more volatile than hours. Chart averages many quarters together.

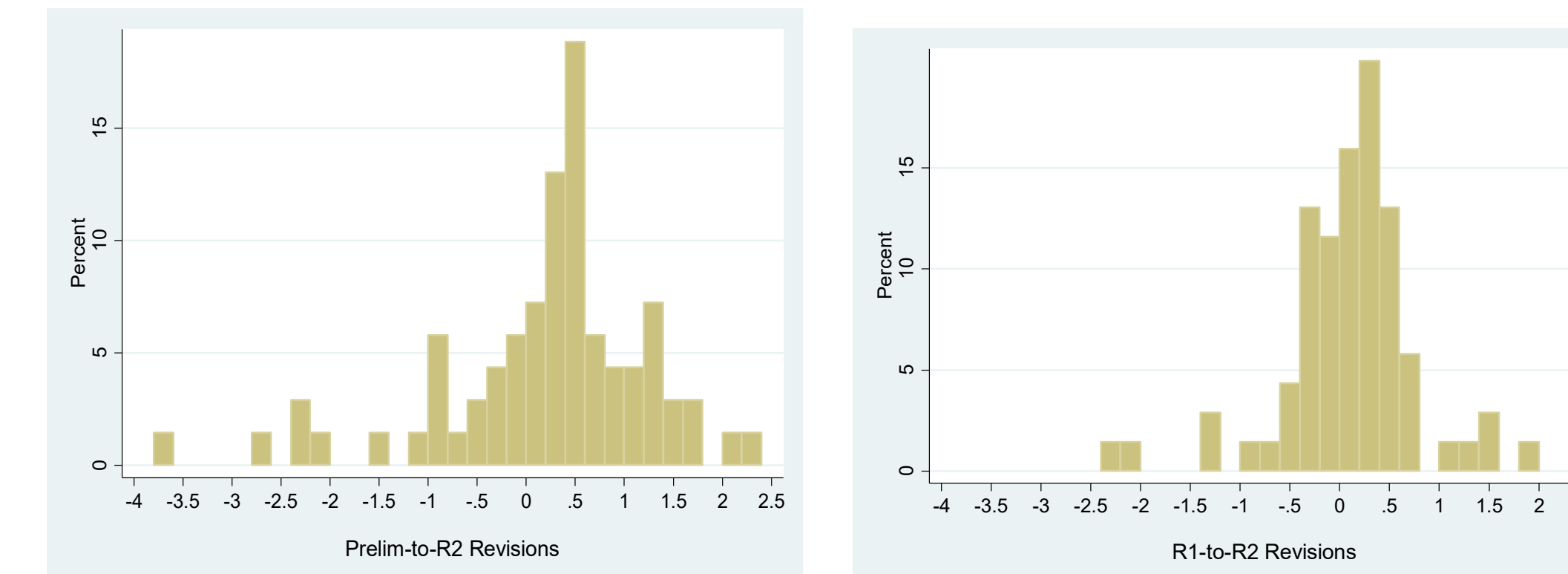
GDP revisions are large as long as new data come in to BEA. The 2013 comprehensive revision raised the level of GDP during 1993-2012 by an average of 3.1% from 1993 to 2012. Mean Absolute Revisions of GDP after comprehensive revisions average about .5. (Fixler et al.)



## Seasonal differences across reference quarters

Output estimates tend to be revised down over time in this period, especially for the first quarter. First quarter hours data, not shown, tends to be revised down also. Revisions aren't generally larger for recession periods, Economic Census years, or other time effects, except for 2008Q4.

## Short run labor productivity revisions before R2



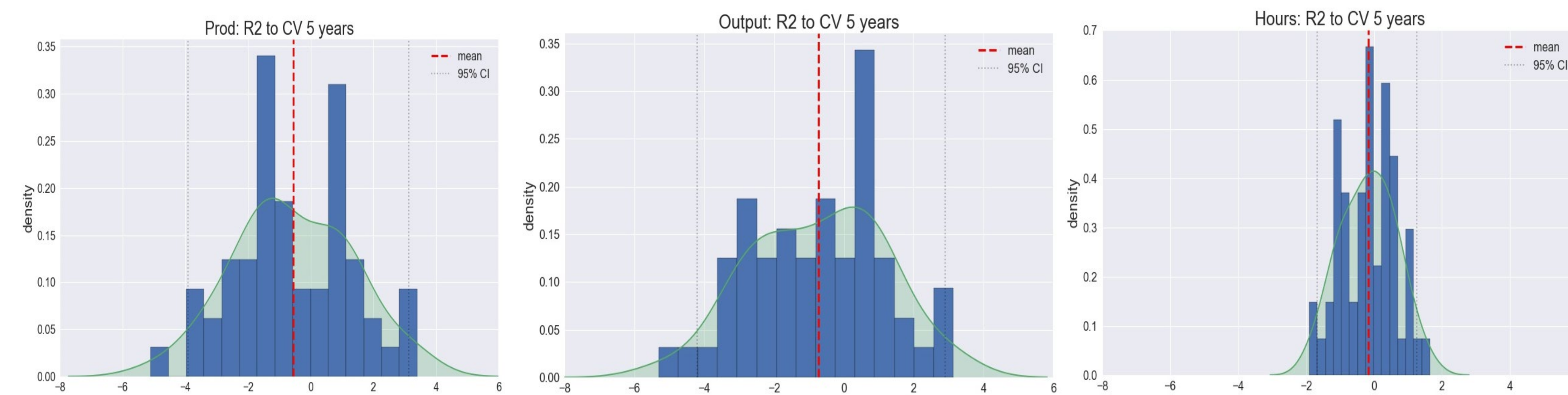
## Distribution of revisions

Revisions have more weight in the tails than a normal distribution. Shapiro-Wilk and other tests reject hypothesis of normal distribution.

2008 Q4 is an extreme outlier. This was the deepest recession quarter. Output was revised sharply down several times.

Revisions in hours especially after R2 have a distribution that is closer to normal.

## Long term revisions after R2, in productivity, output, and hours



## How to characterize short run prediction intervals?

In news releases we characterize ranges of prediction intervals to the third regular release, R2, and do not comment on later revisions. We are reevaluating how to measure and present this information.

Possible methods of using this history to publish updated prediction intervals in news releases:

- Modified standard confidence interval (like BEA)
- Root mean squared error (RMSE) intervals
- Regression-based intervals
- Simple or weighted percentile-based intervals (like the Fed)

Likely ranges of percentiles – 70% range and 90% range. We compared to the practices of other institutions including the Fed and BEA.

For long term (5 year) revisions we are doing new research.

## Methods to make 70% or 90% prediction intervals, compared by cross-validation

Method	Lower*	Upper*	Interval Width*	Percent in Interval
Modified CI	-1.00	1.31	2.3	72.5
RMSE	-1.16	1.16	2.33	73.9
Regression-Based w/Q dummies	-1.03	1.34	2.37	76.8
Percentile Simple	-0.90	1.27	2.17	72.5
Percentile Weighted	-0.90	1.23	2.13	68.1

\* Average over all prediction intervals.

Method	Lower*	Upper*	Interval Width*	Percent in Interval
Modified CI	-1.68	1.99	3.68	89.9
RMSE	-1.86	1.86	3.71	89.9
Regression-Based w/Q dummies	-1.73	2.05	3.79	89.9
Percentile Simple	-2.4	1.7	4.1	92.8
Percentile Weighted	-2.36	1.61	3.97	89.9

\* Average over all prediction intervals.